

White Paper Report

Report ID: 104109

Application Number: HD5135711

Project Director: Leilah Lyons (llyons@uic.edu)

Institution: University of Illinois at Chicago

Reporting Period: 9/1/2011-8/31/2012

Report Due: 11/30/2012

Date Submitted: 2/8/2013

White Paper

Grant # 20110223-HD

CoCensus: Collaboration Exploration of Census Data in a Museum

Leilah Lyons

University of Illinois at Chicago

Submitted 02/03/13

Narrative Description

A. Project Activities

We sought the Level II Digital Humanities Start-Up Grant of \$49,999 to formatively prototype and pilot-test an innovative museum exhibit design that would allow museum visitors to collaboratively explore United States Census data by interacting with a dynamic data map display. Large data sets like the census have traditionally been examined mostly by researchers – not the general public –and in individual research settings – not in public. We employed technology that allowed us to create an exhibit that, although shared by visitors, is still able to respond to each visitor uniquely, allowing visitors to co-investigate a large Census data set in a new way, through the lens of their individual demographic identities. Our institutional partner, the Jane Addams Hull-House Museum, has a long history of using census data to effect social change – the 1895 publication of the *Hull House Maps and Papers: A Presentation of Nationalities and Wages in a Congested District in Chicago, Together with Comments and Essays on Problems Growing Out of the Social Conditions* used groundbreaking illustrations of census data to successfully advocate for social changes in an urban immigrant community.

Our aim with this project was to build on this tradition by creating a 21st century version of such maps. This project addressed two priorities of the Digital Humanities Start-Up Grant: we wished to make complex spatial-historic digital resources like Census data more accessible to the general public, and we wished to situate that access within a public educational context that has the appropriate supports (both ancillary content and docents) to encourage people to connect “dry” demographic data to their own personal and familial histories.

This work was conducted in four distinct strands: Content development, Interaction development, Visualization development, and Funding and dissemination. This report frames our activities in terms of our original objectives, noting instances where tasks were accomplished more smoothly, and where tasks were delayed. Overall, the project met most of the planned objectives.

Strand I: Content Development

Period	Planned Goals:	Attained?
Fall 2011	Develop educational objectives building on visitor interests re: migration histories (Radinsky, Roberts, Lee, Junkin)	Partially: defined pedagogical strategies appropriate for Hull House. Completed May-June 2012.
Spring 2012	Focus group assessment of educational experience <i>in situ</i> in museum (Radinsky, Roberts, Lee, Junkin)	Partially: conducted series of open-ended formative trials in lieu of focus group. Completed during summer of 2012.

To initiate the content development process, we decided that our technology and design team first needed to better understand the context of the Hull-House visit (both self-guided and docent-led). Several members of our team (Lyons, Cafaro, and Roberts) attended the museum first in a self-guided fashion, and later as part of a docent-led tour. This helped us understand the content already being presented at the Hull-House, the style of that presentation of content, and the manner in which docents helped visitors interpret that content. We were able to identify a number of core pedagogical strategies held by the Hull-House institution, which we have incorporated these pedagogical strategies into our thinking about the design of the exhibit experience.

We subsequently had a series of exchanges with our Hull-House colleagues, including an initial demo of technology in our lab (attended by Junkin) and two multi-hour meetings (attended by all grant personnel) to brainstorm content programming ideas surrounding the technology. While this process generated a number of options for the content, it was not as fruitful as we had originally

hoped for, partially (we believe) because the incomplete nature of the technology made it hard for our partners to envision how a completed exhibit would look and feel, and because in early meetings there were too many open questions about which historical Census data would be needed to support different ideas for exhibit development (initial design work proceeded using Census 2000 data only). We were nonetheless able to complete two formative studies that yielded insights on visitor perceptions of Census ancestry categories, and how visitors did or didn't bring their own personal stories of place or family to bear when interacting with the exhibit.

The planned focus group never occurred because we never reached the step of defining a firm set of educational objectives for the exhibit – the mismatch between what we had available in terms of digital census data and the kinds of stories the Hull-House wanted to tell was too large. In the absence of a firm set of educational goals for a focus group to review, we opted to conduct a series of open-ended *in situ* formative trials with Census 2000 mapped first-ancestry data to elicit immigration-related discussions from the visitors themselves, as they posed questions and expressed interests in possible data to explore while using the exhibit. An analysis of a pilot study was presented at the American Educational Research Association's 2012 annual conference (see Appendix). Analysis of the full set of trials, conducted over Summer 2012, is ongoing.

Strand II: Interaction Development

Period	Planned Goals:	Attained?
Fall 2011	Refinement of exhibit prototype's physics model in Dr. Lyons' lab (<i>Lyons, Cafaro</i>)	Yes. Performed formative test to select parameters for physics-based interactions
Spring 2012	Refinement of exhibit prototype's interaction design <i>in situ</i> in museum (<i>Lyons, Cafaro</i>)	Yes. Initial tests validated use of proximity-fading, and suggested replacing jumping with gestures (this work is ongoing). Additional work to seamlessly identify visitors completed.

Our Fall 2011 Interaction development goal was to refine the physics model guiding the response of the visualized data to visitor movements. This work was completed by RA Cafaro and a CS Masters student, Vikrem Bhagi, recruited to join the project (Mr. Bhagi was not listed as a member of the research personnel because he was not financially compensated, but instead used his work as the basis for his Master's Project). In the process of refining the computer algorithms controlling the response of the visualized data to visitors' embodied actions, we generated a theoretical framing for embodied data interaction. We were able to identify several interaction design features we could manipulate (*temporal synchronicity*, *degree of sensitivity*, and *animation "realism"*) to impact visitors' interpretation of the visualization. We selected these three features because they resulted in different tradeoffs with respect to (a) the perceived responsiveness of the system, and (b) the representational clarity of the visualization, and via *in situ* experimentation discovered settings that would be optimal for each of these design features.

We also discovered that visitors felt awkward about jumping or crouching to move their data points, which has fueled a new plan to use gesture recognition to allow visitors to interact with their data sets (e.g., raising their arms above their shoulders to "lift" data points). This iterative refinement was planned for Spring 2012, but was postponed until December 2012 to allow RA Cafaro to rewrite the motion-recognition code using the newly-released Kinect for Windows SDK (Software Development Kit), a new Application Programming Interface (API) that offered improved responsiveness over our old API (Primesense OpenNI). In addition, RA Cafaro was able to perform some previously-unplanned work to streamline the detection and identity assignment of visitors, using a Bayesian regression technique borrowed from Artificial Intelligence that allows us to merge the sensor input from the Kinect camera and the sensor input from our RFID system. This approach allows us to combine the best features of the Kinect

camera (highly precise tracking) with the best features of the RFID system (reliably unique identification of visitors) so the system can dynamically add and remove visitor data from the display as visitors enter and exit, and won't confuse visitors for other visitors in the space. This merged sensing system was validated with an experimental trial *in situ* in early March 2012.

Strand III: Visualization Development

Period	Planned Goals:	Attained?
Fall 2011	Creation of <i>Processing</i> -based visualization (Lyons, Radinsky, Cafaro, Roberts, Beveridge)	Yes. We created the visualization code that renders the visualization and the initial visualization design. We later revised the system to use OpenGL to render the data, as we found <i>Processing</i> to be inefficient.
Spring 2012	Visualization feature set revised <i>in situ</i> in museum in consultation with docents (Radinsky, Roberts, Beveridge)	Yes. Initial visualization was evaluated <i>in situ</i> , with alternate design iterations tested over the summer of 2012.

The schedule called for our team to use the *Processing* programming language in Fall 2011 to create a visualization platform to support the easy importation of GIS-based census data and to respond to user actions. The initial visualization code was accomplished according to schedule, but later design iterations were delayed. Roberts and co-PI Radinsky were able to analyze an initial formative trial of alternative visualization designs with Hull-House visitors, presented at the 2012 International Conference of the Learning Sciences, the largest Learning Sciences conference (see Appendix). RA Cafaro also completed a complete rewrite of the visualization code over the summer of 2012 to use the OpenGL graphics library directly in lieu of using *Processing*. We found that *Processing*, although easy to create and iterate with, did not offer the responsiveness we needed when we began testing the system with multiple visitors.

RA Roberts used the findings presented in the ICLS and AERA papers to inform a new round of visualization design, which was tested in the Summer of 2012. Some changes related to the types of map elements to include in the visualization (e.g., features like parks and transit lines, and labels for specific roads and freeways that visitors commonly referenced), while others related to the scaling of the data itself (i.e., how the size of the centroids is scaled as the size of populations increase across Census tracts). In particular, we found that visitors had a strong tendency to make up narratives to explain the patterns they saw in the data; we used the Summer 2012 *in situ* formative assessments to gather the kinds of narratives visitors are interested in exploring, as well as a list of “misconceived narratives”: commonly-shared narratives that may conflict with evidence observable in the census data. The initial analysis of this *in situ* trial was submitted to the Computer-Supported Collaborative Learning (CSCL) 2013 conference, and is ongoing.

Strand IV: Funding and Dissemination

Period	Planned Goals:	Attained?
Fall 2011	Lyons & Radinsky apply for IMLS <i>Museums for America</i> grant	No.
Spring 2012	Lyons & Radinsky apply for IMLS <i>National Leadership</i> Grant	No. In lieu of pursuing this grant we sought, and won, an ~\$800,000 NSF INSPIRE award, which is a new multi-disciplinary grant created in early 2012.

The plans to apply for the IMLS *Museums for America* and *National Leadership* grants were put on hold, as both of these grants are intended to support the full-scale implementation of a mature project, which our slow Strand I development precluded. Moreover, our preliminary work indicated that there is still a rich amount of research that can (and should) be performed on the

design of “Big Data” interactives, which neither of these grant opportunities would support. In January of 2012, PI Lyons accepted a partial appointment as the Director of Digital learning at the New York Hall of Science, which opened an opportunity to collaborate more closely with project advisor Dr. Andrew Beveridge (Sociology), who is at the nearby CUNY Queens, and is an innovator in the geographic representation of social data. We sought out and won another option for further funding – an NSF INSPIRE interdisciplinary grant. Further details on this Strand are presented in Accomplishments, Continuation of the Project, and Grant Products sections of this document, below.

B. Accomplishments

Despite the delays with Strands I (Content) and III (Visualization), we were able to accomplish the bulk of our goals for the project (see highlighted goals in each of the strands presented in Section A. Project Activities). We completed a fully-functional demo of *CoCensus*, and met our goals for publication and dissemination, reaching all of our target audiences (Education, Learning Sciences, and Human-Computer Interaction researchers; and informal learning practitioners. See Grant Products section for details). We exceeded our goals for further funding, parlaying the work conducted under this NEH grant into a larger, more research-oriented NSF INSPIRE grant which will allow us to research in more depth the issues we uncovered during our initial development. One goal that we did not achieve was creating an exhibit experience that the Hull-House Museum would continue to host. As detailed in the Content Strand of our Project Activities, we were unable to come to a consensus on the *CoCensus* learning goals (pun intended). Limitations of available digital data sets made it hard for the Hull-House staff to imagine how to incorporate *CoCensus* into their interpretive activities, although they remained extremely gracious about allowing us to continue *in situ* testing of the exhibit. We were able to conduct two lab-based experiments and a series of formative *in situ* evaluations of the *CoCensus* exhibit over the course of this grant, addressing a range of design issues from identifying content interests of visitors, eliciting visitor perceptions of map features and data visualizations, and obtaining feedback from visitors on the interaction design.

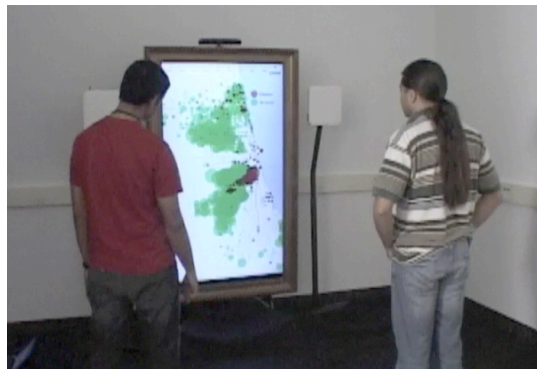


Figure 1. *CoCensus* demo exhibit installed in the Learning Sciences Research Institute (LSRI) lab space on the UIC campus.

C. Audiences

We conducted our formative *in situ* evaluations with typical Hull-House Museum visitors, who tended to be adults who had a personal or professional interest in the themes of the Hull-House Museum (e.g., immigration reform, public health, education, civic engagement). These visitors tended to be highly educated (at least some college, if not graduate study), and came from a wide variety of geographic locales (all over the United States, as well as from other countries). Owing to the decision not to install a permanent version of *CoCensus*, we did not come into contact with

any visitors outside those who participated in our studies (around 90 individuals). We cannot make any claim that *CoCensus* drove attendance – most visitors were surprised to see a digital exhibit at the Hull-House. Our hope had been to create a permanent exhibit that would attract people from current immigrant communities within Chicago, but achieving that would have required some outreach and programmatic coordination. The Hull-House regularly hosts events that draw members of the local communities, but none of these occurred during our testing sessions so we were unable to gauge whether this local demographic would find *CoCensus* engaging or not. We maintain friendly and collaborative communications with our Hull House partners, and hope to revisit the exhibit ideas with them when the current version of the exhibit has been completed, which may allow for clearer ideas for designs that would better suit Hull House's mission and vision.

Our current NSF INSPIRE grant is supporting the installation of *CoCensus* at the New York Hall of Science (NySci), located in Queens, New York. NySci serves one of the most diverse visitor audiences in the United States, and skews quite a bit younger than the Hull-House, given that its main demographic is visiting school groups. We anticipate that over the lifetime of the INSPIRE grant *CoCensus* will reach tens of thousands of visitors at NySci.

D. Evaluation

This project was not externally evaluated – we conducted certain kinds of evaluation (e.g., of the visualization and interaction designs) via formative trials with around 90 visitors, the results of which are detailed in the papers included in the Appendices. In addition, Cafaro and Bhagi conducted an experimental formative trial *in situ* at the Hull-House Museum with 9 users, where each user experienced each of the 12 different interaction conditions (2x2x3) in a randomized order and provided subjective ratings. We found that a particular combination of interaction design features was highly preferred by the users (interaction which is continuous in its response to visitor movements, as opposed to being "triggered" by specific motions; interaction which is asynchronous in its response, meaning that there is a lag between visitor actions and the system's response; interaction which animates the data using a "true physics" model, as opposed to a cartoon-like exaggerated physics model or a model which responds linearly to visitor motion).



Figure 2. On the left, a photo of an initial trial of CoCensus in the Maps and Papers exhibit at the Hull-House museum, and on the right, a photo of the experimental formative trial that elicited visitors' preferred interaction design features. Experimentors used large colored magnets to mark a visitor's preferences along Likert scales posted on an easel as they experienced the 12 conditions.

The main weakness of the project was the difficulty bridging between our current designs of interactive mapped census data, and the Hull House partners' desired themes exploring identity and civic efficacy. Together we explored several possible immigration themes that would suit the Hull-House, like the impact of the 1917 and 1924 Immigration Acts on the demographic composition of Chicago neighborhoods (in particular, how the 1924 quota system affected Southern and Eastern European immigration). It was thought that this theme would be a good fit, given the large Italian immigrant community near the Hull-House Museum (persisting to this day) but we found that the Census data available to us was not well suited to allowing visitors to explore this era of immigration history in Chicago. There were difficulties presented by changing ancestry, nativity, ethnicity, and race categories over the history of the U.S. Census. These changes pose valuable questions for investigating how identity categories are mediated by government decisions, but make it difficult to visualize changes in a given variable earlier than 1990. A bigger problem was that most Census data variables are not available at the census-tract or block-group geographic level for these early time periods. For example, even though Census tracts were defined for the first time in Chicago in 1910, we could not obtain data at a tract-level granularity via the National Historic Geographic Information System distribution site (www.NHGIS.org). This made pre- and post-Immigration Act comparisons of Chicago demographic distributions difficult: the only data we could locate was at the county level, which relegated all of Cook County (subsuming Chicago) to being a single data point. The 1893 Hull-House Maps and Papers project, the original inspiration behind this proposal, collected data on a house-by-house basis, but we could not obtain digital versions of this information, and did not have the time or resources to compile new digital versions at this granularity from the original Census archives. Our Hull-House partners felt that local-level granularity (e.g., block level data) was needed to complement their interpretive goals. Our formative evaluations confirmed this – visitors were highly interested in connecting demographics to discrete, recognizable places on the map of Chicago (e.g., neighborhoods, landmarks). While block-level data *are* available for the 2000 and 2010 census surveys, we could not find an immigration-related theme in this time frame that our Hull-House partners felt would complement the existing Maps and Papers exhibit. When we ran into difficulty obtaining the originally-envisioned data sets to use in the exhibit, it was hard for us to find a substitute exhibit theme that Hull-House staff felt they could work into their existing interpretive programming.

The UIC team lacked experience with designing docent facilitation, and so wasn't able to help bridge this gap – we could suggest broad ideas (e.g., changes in neighborhood demographics) but had a hard time phrasing these ideas in a way that helped the Hull-House staff see how they might be used to engage visitors. This gap in our communication – the development team's difficulty conveying the potential affordances of the exhibit, and the Hull-House staff's difficulty helping us understand how to approach interpretation – meant our brainstorming sessions did not produce a design that met the educational goals of both partners. To correct this problem, for our new NSF INSPIRE grant we have recruited a Learning Sciences PhD student who has worked as a docent and as a docent trainer at a number of historical and scientific informal learning institutions. Her ability to dive in and design a range of different interpretation scripts after only a few weeks of participation suggests that what we needed was someone with interpretation expertise who could spend significant time “embedded” in project design meetings.

E. Continuation of the Project

Drs. Lyons, Beveridge, and Radinsky applied for and were awarded an \$800,000 NSF INSPIRE (formerly dubbed “CREATIV”) grant, which is allowing us to take a more research-focused tack to exploring the Big Data interaction issues raised by *CoCensus*. For this new grant we will explore how learners come to reason about geospatially-plotted data in three different learning

environments (museums, classrooms, and on the web). We have targeted several learning challenges identified in the literature and by the studies we conducted under this NEH grant. Although the Hull-House has indicated that it does not see a role for CoCensus as a permanent exhibit, the positive relationship we built with the staff resulted in a gracious offer to use their museum as a testing site, much as we did during this grant. We have secured a space at the New York Hall of Science to install a version of *CoCensus* which we will use in further iterative design-based research. We have also developed a partnership with the *Social Explorer* company (of which Beveridge is the founder and Director), which produces a website of the same name that allows users to interactively explore Census data.

F. Long Term Impact

Although this NEH grant did not result in a permanent exhibit at the Hull-House, it has allowed us to springboard to another context (the New York Hall of Science) which will give our work even greater reach, under the auspices of a 3 year NSF INSPIRE grant. Our hope is that the research work conducted both under this NEH grant and the new NSF grant will impact both researchers and practitioners as they grapple with how to present large data sets in informal learning institutions. We hope to return to our Hull House partners in the future to revisit the idea for which we share enthusiasm, once we have examples that can be used productively as a springboard for design.

F. Grant Products

A large amount (over 40,000 lines) of programming code was developed for this project (code to display georeferenced Census data, code to interpret and merge the sensor inputs from the Kinect camera and RFID subsystems, code to control visualization responses to sensor inputs). It is still under development but will be made publicly available when closer to completion. This grant resulted in 5 academic publications (see Appendix). The first formative analysis was written up in the paper Roberts presented at the AERA 2012 conference (the largest educational research conference, sponsored by the American Educational Research Association), and insights from this analysis were condensed into design guidelines in the form of a short poster paper presented at the International Conference of the Learning Sciences (ICLS) 2012 conference (the largest Learning Sciences conference). The second formative study, conducted in summer of 2012, examined how the embodied interaction affected visitors' perceptions of the data being displayed and their own personal immigration stories. An analysis of this study was submitted to the Computer-Supported Collaborative Learning Conference (CSCL), the "sister" conference of ICLS, as a full paper (review pending; see Appendix). RA Cafaro's work developing a sensor-merging system was written up as a full paper (see Appendix) accepted to the CHI 2013 conference (an Association of Computer Machinery-sponsored conference devoted to Human-Computer Interaction; the premiere conference in the field). RA Cafaro was also able to parlay work performed on this project into an idea for his dissertation research, which attempts to identify how to unify interaction metaphors as suites of "design allegories" for embodied interaction, work which was accepted into the Doctoral Consortium of the most highly-regarded non-desktop human-computer interaction conference, UBICOMP (see Appendix).

In terms of reaching out to the practitioner community, we produced and presented a demo video of *CoCensus* (<http://www.youtube.com/watch?v=MIE2hliOSu0>) at the invitation of the Association of Science and Technology Centers (ASTC) 2012 conference, where Dr. Lyons was also asked to serve on a keynote panel on Cyberlearning in informal learning institutions. As a result of her participation in the panel, Dr. Lyons was asked to pen an article for ASTC's magazine *Dimensions* on her work with computer-based exhibits like *CoCensus*. This magazine

is targeted at informal learning institute practitioners.

Appendix Contents

Included in this Appendix are five papers, four accepted to the AERA 2012, ICLS 2012, UBICOMP 2012, and CHI 2013 conferences, and one submitted to the CSCL 2013 conference (still under review). The citations would be:

Cafaro, F. (2012). *Using embodied allegories to design gesture suites for human-data interaction*. Paper presented at the Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UBICOMP '12).

Cafaro, F., Panella, A., Lyons, L., Radinsky, J., & Roberts, J. "I See *You* There! Developing Identity-Preserving Embodied Interaction for Museum Exhibits," Accepted to the *ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*, Paris, France, 2013.

Roberts, J., Radinsky, J., Lyons, L., & Cafaro, F. "Co-Census: Designing an Interactive Museum Space to Prompt Negotiated Narratives of Ethnicity, Community, and Identity," Presented at *American Education Research Association (AERA '12)*, Vancouver, CA, 2012.

Roberts, J., Lyons, L., & Radinsky, J. "Connecting Visitors to Exhibits through Design: Exploring United States census data with CoCensus," In the *Proceedings of the 10th International Conference of the Learning Sciences (ICLS '12)*, Sydney, Australia, ISLS, 2012.

Roberts, J., Lyons, L., Radinsky, J., & Vogt, K. (submitted for review). "That's Me and That's You: Museum visitors' perspective-taking around an embodied interaction data map display." Paper submitted to the *Computer-Supported Collaborative Learning Conference (CSCL '13)*, Madison, WI, ISLS, 2013.